Redes Neuronales y Aprendizaje profundo

Tema 0 – Repaso de Álgebra Lineal

Irina Arévalo

Contenido

1. Operaciones con matrices - vectores

- Los vectores son filas o columnas de datos. Normalmente se representan por letras minúsculas.

- En análisis de datos, los vectores son columnas de *n* datos es decir, con dimensión *n x 1.*

- Un vector *3 x 1* es:

$$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

- Los vectores tienen una representación gráfica y un vector *n x 1* es un punto en un espacio de dimensión *n*

1. Operaciones con matrices - matrices

- Una matriz es una disposición de números en forma de rectángulo. Las matrices se denotarán con letras mayúsculas en negrita **A, B,…**

- La dimensión de una matriz es el número de filas y columnas, y se escribe $f \times c$. Los elementos de una matriz se identifican por su posición en la fila y la columna.

- La transpuesta de una matriz **A** con dimensión $f \times c$ es el resultado de cambiar en **A** las filas por las columnas, y se denota por $\mathbf{A}^t$, que tiene dimensión $c \times f$. Si $\mathbf{A} = \{a_{ij}\}$ entonces $\mathbf{A}^t = \{a_{ji}\}$

- La transpuesta de un vector columna es un vector fila y viceversa.

1. Operaciones con matrices - matrices

- La matriz 3×3

$$\boldsymbol{A} = \begin{pmatrix} 3 & 5 & 8 \\ 5 & 7 & 1 \\ -4 & 8 & -3 \end{pmatrix}$$

Se representa en Python como:

```
import numpy as np
A = np.array([[3, 5, 8], [5, 7, 1], [-4, 8, -3]])
```

1. Operaciones con matrices

- Las matrices pueden sumarse y restarse únicamente si son de la misma dimensión

$$\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij}), \quad \mathbf{A} - \mathbf{B} = (a_{ij} - b_{ij})$$

- Producto de una matriz **A** por un escalar λ

$$\lambda\mathbf{A} = (\lambda a_{ij})$$

- Las matrices pueden mutiplicarse sólo si las dimensiones son compatibles. Si **A** tiene dimensión *f x c* y **B** tiene dimensión *c x d* entonces **AB** tiene dimensión *f x d*.

- Las matrices no se pueden dividir. Si se da la compatibilidad de dimensiones, entonces se puede obtener el producto de **A** por $\mathbf{B}^{-1}$ si existe la inversa de B.

1. Producto de matrices

- El producto de dos matrices es otra matriz donde cada elemento es el producto de fila por columna.

- Si **A** = $\{a_{ij}\}$ es *m x n* y **B** = $\{b_{ij}\}$ es *n x p*, entonces **C** = **AB** es *m x p*, donde **C** = $\{c_{ij}\}$ = $\left\{\sum_{r=1}^{n} a_{ir} b_{rj}\right\}$

- En general **AB** $\neq$ **BA**, salvo para algunas matrices especiales.

1. Matrices especiales

Las siguientes matrices son todas cuadradas:

- Matrices diagonales: únicamente los elementos de la diagonal pueden ser diferentes de cero
- Matrices simétricas: $\mathbf{A} = \mathbf{A}^t$
- Matrices invertibles o matrices no singulares, son aquellas que $\mathbf{A}^{-1}$ existe
- Matrices ortogonales: son aquellas que verifican: $\mathbf{A}^{-1} = \mathbf{A}^t$
- Matrices triangulares superiormente (inferiormente): son aquellas donde únicamante los elementos por encima de la diagonal (debajo) son distintos de cero.

1. Matriz identidad

- La matriz identidad, **I,** es una matriz diagonal que contiene sólo unos

- La matriz identidad *2 x 2* es:

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

- La matriz identidad *3 x 3* es:

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- Para cualquier matriz cuadrada A:

AI = IA = A

Contenido

1. Operaciones con matrices
2. Dependencia y rango en matrices
3. Autovalores y autovectores
4. Descomposición y diagonalización

2. Dependencia lineal

- La *dependencia lineal* se refiere al caso donde existe una relación lineal entre filas o columnas de una matriz.

- Cuando existe dependencia lineal, se dice que tenemos una matriz *singular*, porque no tiene inversa (matrices cuadradas).

- En general, el número de filas (o columnas) linealmente independientes se llama *rango* de la matriz.

- Si todas las filas (o columnas) son linealmente independientes, se dice que la matriz tiene *rango completo.*

- Las matrices de rango completo son *no singulares.*

2. Rango de una matriz

- En Python obtenemos el rango de una matriz con

```
from numpy.linalg import matrix_rank
matrix_rank(A)
```

2. Inversa de una matriz

- La inversa de una matriz **A** es la matriz **A^{-1}** tal que

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

- Por tanto, el producto de una matriz por su inversa es la matriz identidad
- No todas las matrices tienen inversa, pero si existe es única
- Únicamente las matrices no singulares tienen inversa
- Un test rápido para saber que una matriz es no singular (es decir, es invertible) es probar que su determinante no es cero.

2. Determinante de una matriz

- Para una matriz cuadrada *2 x 2* **A** tal que

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

su determinante es la cantidad *ad – bc.*

- Se suele representar como

$$\det(\mathbf{A}) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

- El determinante de una matriz 3 x 3 es

$$\begin{vmatrix} a & b & c \\ u & v & w \\ x & y & z \end{vmatrix} = a \begin{vmatrix} v & w \\ y & z \end{vmatrix} - b \begin{vmatrix} u & w \\ x & z \end{vmatrix} + c \begin{vmatrix} u & v \\ x & y \end{vmatrix}$$

2. Propiedades de los determinantes

- Los determinantes sólo existen para matrices cuadradas.
- Las siguientes propiedades son útiles para el cálculo de determinantes:
  - Para cualquier matriz cuadrada: $|\mathbf{A}^t| = |\mathbf{A}|$
  - Para cualquier matriz invertible: $|\mathbf{A}^{-1}| = |\mathbf{A}|^{-1} = \frac{1}{|\mathbf{A}|}$
  - Para una constante c y una matriz $\mathbf{A}$ $n \times n$: $|c\mathbf{A}| = c^n|\mathbf{A}|$
  - Si $\mathbf{A}$ y $\mathbf{B}$ son matrices $n \times n$: $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$

2. Operaciones con matrices en NumPy

Operador	Descripción
array()	Crea una matriz
dot()	Multiplicación de matrices
transpose()	Matriz traspuesta
linalg.inv()	Matriz inversa
linalg.det()	Determinante
flatten()	Transforma una matriz en un array de 1D
linalg.solve(A, y)	Resuelve el sistema Ax = y

Contenido

3. Autovalores y autovectores

- Desde el punto de vista geométrico, una matriz cuadrada *n x n* puede considerarse como una transformación en $\mathbb{R}^n$.

- Si al vector *v* = $(1, 3)^t$ le aplicamos una matriz obtenemos:

$$\mathbf{A}v = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} = 2v$$

- Por tanto **A** traslada el vector *v* a otro vector que es un múltiplo de si mismo, es decir, se encuentra en la misma dirección.

- El vector *v* se denomina *vector propio o autovector* de **A** (puesto que se transforman en un múltiplo de ellos mismo por **A**) y la constante 2 *autovalor* de **A** asociado a *v*.

3. Autovalores y autovectores

- De manera formal, si **A** es una matriz *n x n*
  1. Cualquier vector x tal que $\mathbf{A}x = \lambda x$ para alguna constante λ se denomina *autovector* de **A**
  2. El valor λ es el *autovalor* asociado a *x*

- En otras palabras:
  1. Los autovectores son vectores que son trasladados por la matriz a otro vector en la misma dirección a través del origen
  2. El correspondiente autovalor de un autovector es un *factor de escala*.

3. Autovalores y autovectores

- Si x es un autovector, entonces kx es también un autovector con el mismo autovalor

$$\mathbf{A}x = \lambda x; \qquad \mathbf{A}(kx) = \lambda(kx)$$

- Para evitar esta indeterminación, en ocasiones, los autovectores se normalizan de modo que $|x| = 1$. Sin embargo el signo no está determinado: si x es un vector propio, también los es $-x$.

3. Autovalores y autovectores

- En la práctica buscamos los x y λ tales que

$$\mathbf{A}x = \lambda x \Rightarrow (\mathbf{A} - \lambda\mathbf{I})x = 0$$

donde **I** es la matriz identidad *n x n*

- Se trata de un sistema homogéneo que tendrá solución no nula si el determinante de la matriz es cero:

$$|\mathbf{A} - \lambda\mathbf{I}| = 0$$

- A la expresión $|\mathbf{A} - \lambda\mathbf{I}| = 0$ se le denomina *polinomio característico*.

3. Autovalores y autovectores

Ejemplo

- Encontrar los autovalores de la matriz $A = \begin{pmatrix} 2 & -1 \\ -2 & 3 \end{pmatrix}$
- El polinomio característico es:

$$\begin{vmatrix} 2-\lambda & -1 \\ -2 & 3-\lambda \end{vmatrix} = (2-\lambda)(3-\lambda) - 2 = 0$$

- Las soluciones de la ecuación $\lambda^2 - 5\lambda + 4 = 0$ son $\lambda_1 = 1$ y $\lambda_2 = 4$, que son los autovalores de A. Los autovectores son los vectores $v = (x, y)^t$ tales que $(A - \lambda_1)v = 0$ y $(A - \lambda_2)v = 0$

- En Python usamos el comando `linalg.eig(A)`

3. Propiedades

- Una matriz **A** *n x n* tiene *n* autovalores. Sin embargo, algunos de ellos pueden ser cero, otros reales (y algunos repetidos) y pueden ser también números complejos.

- Si *x* es un autovector de **A**, *cx* es también un autovector de **A**.

- Si la matriz **A** es diagonal, sus autovalores son los elementos de la diagonal.

- Si la matriz **A** es simétrica, entonces los autovalores son números reales.

- Si **A** es simétrica, sus autovectores son ortogonales. Es decir, dos autovectores con diferente autovector tienen un producto escalar que es igual a cero.

3. Propiedades

- La *suma de los autovalores* de **A** es igual a la traza de **A**:

$$\text{traza}(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i$$

- El *producto de los autovalores* de **A** es igual al determinante de **A**:

$$|A| = \prod_{i=1}^{n} \lambda_i$$

- El rango de **A** es el número de autovalores no nulos.

- Si **B** es una matriz no singular, las matrices **A** y $\mathbf{B}^{-1}\mathbf{AB}$ tienen los mismos autovalores.

Contenido

4. Descomposición y diagonalización

- *Descomponer una matriz* **A** es escribirla como producto de otras dos: **A = BC**

- *Diagonalizar una matriz* **A** es escribirla como producto de tres matrices de la forma **A = BDC** donde **D** es una matriz diagonal y las matrices **B** y **C** son ortogonales.

- Veremos tres resultados:
    - Descomposición de Cholesky
    - Teorema de descomposición espectral
    - Descomposición en valores singulares SVD

4. Descomposición y diagonalización

- **Descomposición de Cholesky**. Sea **A** una matriz simétrica semi definida positiva. La matriz de Cholesky asociada a la matriz **A** es una matriz triangular inferior que verifica

$$\mathbf{A} = \mathbf{C}\mathbf{C}^t$$

- La matriz **C** se puede interpretar como la raíz cuadrada de **A**
- Al igual que una raíz cuadrada, sólo existe si la matriz es semidefinida positiva
- La matriz no siempre es única

4. Descomposición y diagonalización

- El siguiente resultado de diagonalización se conoce como **teorema de descomposición espectral**.

- Cualquier matriz <u>simétrica</u> **A** se puede escribir como

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^t$$

donde **V** es la matriz ortogonal de los autovectores v_i normalizados y

$$\mathbf{D} = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$$

es una matriz diagonal con los autovalores.

4. Descomposición y diagonalización

- Por último veremos la **descomposición en Valores Singulares o SVD**
- Se refiere a un resultado de diagonalización para matrices no cuadradas
- Si **A** es $m \times n$ y real, los autovalores de $\mathbf{AA}^t$ son siempre reales y mayores o iguales que cero.
- Representamos dichos autovalores por

$$\lambda_1 \geq \ldots \geq \lambda_m \geq 0$$

- A los valores $\tau_i = \sqrt{\lambda_i}$ se les denomina valores singulares de **A**.

4. Descomposición y diagonalización

- **Descomposición en Valores Singulares SVD**. Si **A** es una matriz real *m x n*, admite una descomposición del tipo:

$$\mathbf{A} = \mathbf{U\Sigma V}^t$$

donde **U** es *m x m* y **V** *n x n* son matrices ortogonales, y **Σ** *m x n* es una matriz donde los elementos diagonales son los valores singulares de **A**

4. Descomposición y diagonalización

Descomposición de matrices en Python

Operador	Descripción
`linalg.cholesky(A)`	Factorizacion de Choleski **A**. Devuelve matriz triangular superior de modo que $\mathbf{B}^t\mathbf{B} = \mathbf{A}$
`linalg.eig(A)`	Teorema de representacion espectral de **A**
`U, S, Vh = np.linalg.svd(A)`	Descomposición de valores singulares SVD de **A**

Redes Neuronales y Aprendizaje profundo

Tema 0 – Repaso de Álgebra Lineal

Irina Arévalo